

DOCUMENT RESUME

ED 399 296

TM 025 611

AUTHOR Johanson, George A.; Gips, Crystal J.
TITLE Paired Comparison Intransitivity: Useful Information
or Nuisance?
PUB DATE Apr 93
NOTE 6p.; Paper presented at the Annual Meeting of the
American Educational Research Association (Atlanta,
GA, April 12-16, 1993).
PUB TYPE Reports - Research/Technical (143) --
Speeches/Conference Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Administrator Attitudes; *Comparative Analysis;
*Likert Scales; *Principals; Responses; *Test Format;
Test Items
IDENTIFIERS Choice Behavior; Circular Triad Scores; *Paired
Comparisons; Ranking; Rank Order

ABSTRACT

The decision to use a forced-choice test item format versus an item format where choice is not forced (e.g., a Likert scale) might best be determined by the nature of the information sought since the difficult decisions required for forced-choice formats may result in a different scaling than an unforced method. If a forced choice is desired, there remains the question of whether to have subjects rank order their preferences or to present the items as paired comparisons. One of the main differences between ranking and paired comparisons is the potential in the latter format for what is known as "circular triads" or "intransivities." A subject may choose object A over B, then B over C, and then, intransitively, C over A. Are these intransivities an unnecessary complication or a useful source of information? An example involving the hiring preferences of principals using a national sample of 203 principals with 36 paired comparisons illustrates, through computation of the total number of circular triads, that knowledge of intransivities adds to knowledge about the real choices of respondents. By excluding the possibility of intransivities, a ranking format in the hiring preferences example would have been likely to yield similar scale values but much less insight into the respondents' processes. (Contains 13 references.) (SLD)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

GEORGE JOHANSON

PAIRED COMPARISON INTRANSITIVITY:
USEFUL INFORMATION OR NUISANCE?

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

A paper presented at the annual meeting of the
American Educational Research Association
Atlanta, GA
April 1993

George A. Johanson & Crystal J. Gips
College of Education
Ohio University

The decision to use a forced-choice item format verses an item format where choice is not forced (e.g., Likert) might best be determined by the nature of the information sought since the difficult decisions required for forced-choice formats may result in a different scaling than from an unforced method (Alwin & Krosnick, 1985). If relative comparisons between options are desired and a forced-choice is adopted, there are still item format options. Having the subjects rank-order their preferences is one option; a second option is to present the objects under investigation as paired comparisons. There are scaling procedures (e.g., Thurstone scaling) available that yield interval-level scales with either of these item formats (Dunn-Rankin, 1983). One of the main differences between ranking and paired comparisons is the potential only in the later format for what are known as 'circular triads' or 'intransitivities'. That is, a subject may choose object (or stimulus) A over object B, object B over object C, and then, intransitively, object C over object A. The total number of such circular triads (TCT) for each subject can be calculated with the formulas in Kendall's text (1970). Are these intransitivities an unnecessary complication or a useful secondary source of information? If the former is more often the case, we would be inclined to a ranking format which precludes intransitivities. If the latter is true, paired comparisons could be the format of choice if we also compute TCT.

THE LITERATURE

There is evidence in the literature that intransitivity may be useful. Hendel (1977) found that intransitivity was related to both the respondents' individual differences and the sets of stimulus objects. That is, if the objects are very close to one another with respect to the construct (e.g., preference) being investigated, it could be difficult to reliably rank them at all. In such cases, you would expect the number of TCT to be relatively large. If the construct and objects are multidimensional, you might also expect higher mean TCT.

With individuals or subjects, there have been several factors studied with respect to TCT. In children, age was

PAIRED COMPARISON INTRANSITIVITY

related to TCT (Riechard, 1990) with younger children having higher TCT. Socioeconomic status was related to TCT (Riechard, 1991) with lower SES groups having more TCT. Aptitude was found to be related to TCT (Sharac, 1976) with lower TCT occurring in those with higher aptitude scores. Sex and race were found to relate to TCT (Hallinan, 1988) while in another study, TCT was unrelated to sex (Riechard, 1991). Tinsley et al. (1984) found that intransitivities increased on a post-test of paired comparison values after values clarification instruction. The implication here is that more knowledge may not necessarily reduce intransitivities. Hendel (1979) found that intransitivity in high school students was somewhat stable and correlated about $+.4$ over a two-year period. Budescu and Weiss (1987) found that transitive subjects remained relatively consistent from gain situations to loss situations and that the preferences were perfectly mirrored. With intransitive subjects, however, the circular triads were not mirrored.

AN EXAMPLE: PRINCIPALS' HIRING PREFERENCES

In a mailed survey of hiring preferences (Johanson & Gips, 1992), paired comparisons were used with a national sample of 271 high school principals to determine the relative desirability of 9 traits or characteristics of teaching candidates. For the 203 principals with complete data on the 36 paired comparisons of traits, the mean number of total circular triads (TCT) per subject was 3.85 ($SD=3.83$). There were 39 subjects with perfectly consistent decisions or zero TCT.

TCT related weakly, but significantly, to the number of years experience as a principal (-0.186 , $p=.004$, one-tail) in this group. The correlation of TCT with age of the principal was not significantly different from zero. This indicated that the more experienced principals had somewhat less difficulty ranking these traits consistently. In general, the mean TCT in this study was similar to those seen in other applications of paired comparison scaling methods (Riechard, 1990).

In an initial free-response item, subjects identified a single most desired trait. Responses were coded as being indicative of either cognitive or affective traits. The mean TCT for those ($N=124$) with complete data in the affective group (3.64) was near the group mean and 2 fewer (statistically significant) than the mean TCT for those ($N=24$) in the cognitive group (5.67). Since affective traits were often chosen in the paired comparisons by both those with a cognitive and affective free-response, this suggested that choices may have been more difficult for those with initial cognitive priorities.

To explore this possibility further, a means was sought to identify a second subgroup of subjects who may have been inclined towards cognitive skills. Since subjects were also asked about the role they thought grade point average (GPA) does and should play in the hiring decision, we were able to identify 50

PAIRED COMPARISON INTRANSITIVITY

principals who thought that GPA should play a greater role than it now does and who had complete data on the 36 paired comparisons. The survey also included a Likert scaling of the original nine qualities. Using only the Likert ratings of this subgroup, the cognitive traits were seen to be virtually the same as in the larger group. That is, it would appear that the affective qualities were most attractive even to this subsample of principals who felt that cognitive skills (in the form of GPA) should play a larger role in the selection process.

While this subgroup was similar to the free-response group in this regard, there were only 7 common subjects and the classifications were relatively independent of each other. However, once again, the mean TCT (5.34) of the 'should>does' subgroup (N=50) was significantly greater than the mean TCT (3.42) of the remaining subjects (N=152). The mean TCT of the 7 common subjects was 7.14.

The distinction between 'choices' and 'preferences' can be important in that the former are observed, typically thought to be driven by the later, and sometimes seen to be intransitive (Bar-Hillel & Margalit, 1988). Our data would indicate that those principals who gave indications of a cognitive preference still tended to make largely affective choices. The additional conflict in these groups was only made apparent by the significant increase in TCT for these groups.

Together, the information gained from a knowledge of TCT gave us considerably more confidence in our conclusion that while affective traits were chosen most often by all subgroups of principals, the choices were considerably more difficult for those principals who indicated that they saw cognitive traits as desirable.

CONCLUSIONS

By excluding the possibility of intransitivities, a ranking format in the hiring preferences example would have likely given us similar scale values, but considerably less insight into process. In particular, there would have been no discernible differences between those principals who gave indications of cognitive preferences in either the free-response or GPA items and those who did not.

Also of interest was the fact that the experience of these principals was reasonably unrelated to all other measures except TCT. Unlike the study by Tinsley (1984), the 'post-measure' of TCT (TCT in the more experienced group) did not show an increase due to the expanded knowledge, but rather a modest decrease. Principals' priorities may be complex, but there was evidence that the experience of being a principal may have helped to clarify the choices.

Formulas for the maximum number of possible triads and approximate confidence intervals for TCT with random response are well known (Kendall, 1970). With knowledge of the number of

PAIRED COMPARISON INTRANSITIVITY

circular triads that could be expected by chance in random responses, it is possible to identify those subjects who may have misunderstood the task, lack the requisite skills or knowledge to perform the task, or who simply responded without thought. That is, TCT can provide valuable information for initial data screening that would also not be available with a ranking format.

In short, measures of intransitivity can be useful in that they provide supplementary information that is not elsewhere available and that may clarify issues regarding both stimuli and individuals. On the other hand, using a paired comparison format without computing TCT may well be less desirable than a ranking format since intransitivities may cloud or even disguise important information.

A microcomputer program (IBM or compatible) that computes the maximum TCT possible and TCT for each subject has been written in standard Pascal and is available free of charge from the first author when the request is accompanied by a formatted diskette and stamped, self-addressed mailer.

PAIRED COMPARISON INTRANSITIVITY

REFERENCES

- Alwin, D. F., & Krosnick, J. A. (1985). The measurement of values in surveys: A comparison of ratings and rankings. Public Opinion Quarterly, 49(4), 535-552.
- Bar-Hillel, M., & Margalit, A. (1988). How vicious are cycles of intransitive choice? Theory and Decision, 24, 119-145.
- Budescu, D. V., & Weiss, W. (1987). Reflection of transitive and intransitive preferences: A test of prospect theory. Organizational Behavior and Human Decision Processes, 39,(2), 184-202.
- Dunn-Rankin, P. (1983). Scaling methods. Hillsdale, NJ: Lawrence Erlbaum.
- Johanson, G. A., & Gips, C. J. (1992). Secondary school hiring practices. The High School Journal.
- Hallinan, M. T. (1988). The effects of individual and structural characteristics on intransitivity in social networks. Social Psychology Quarterly, 51(2), 81-92.
- Hendel, D. D. (1977). Intransitivity on paired-comparisons instruments: The relationship of the total circular triad score to stimulus circular triads. Applied Psychological Measurement, 1(3), 403-411.
- Hendel, D. D. (1979). Paired comparisons intransitivity: Is it relatively stable over time? Educational and Psychological Measurement, 39(4), 779-784.
- Kendall, M. B. (1970). Rank correlation methods (4th ed.). New York, NY: Hafner Press.
- Riechard, D. E. (1990). Paired-comparisons intransitivity related to age of subjects. Educational and Psychological Measurement, 50(1), 105-110.
- Riechard, D. E. (1991). Intransitivity of paired comparisons related to gender and community socioeconomic setting. Journal of Experimental Education, 59(2), 197-203.
- Sharac, J. S., Rounds, J. B., & Davis, R. V. (1976, September). Intransitivity and vocational needs. Paper presented at the Annual Meeting of the American Psychological Association, Washington, D. C.
- Tinsley, H. E., Benton, B. L., & Rollins, J. A. (1984). The effects of values clarification exercises on the value structure of junior high school students. Vocational Guidance Quarterly, 2(3), 160-167.